AD-A117 231 SOUTHERN METHODIST UNIV DALLAS TEX DEPT OF STATISTICS F/G 12/1
RELATED CORRELATION COEFFICIENTS.(U)
MAR 82 J E BOYER, A D PALACHEK, W R SCHUCANY N00014-82-K-0207
UNCLASSIFIED TR-158 NL

1 of 1
AD A
117231

END
DATE
FILMED
8-8
DTIC

AD A117231

12

RELATED CORRELATION COEFFICIENTS

by

J.E. Boyer, Jr., A.D. Palachek and W.R. Schucany

Technical Report No. 158
Department of Statistics ONR Contract

March 1982

Research sponsored by the Office of Naval Research
Contract N00014-82-K-0207

DEPARTMENT OF STATISTICS
Southern Methodist University
Dallas, Texas 75275

Related Correlation Coefficients

Key Words: Correlation coefficient, normal scores hypothesis testing, approximately distribution-free

ABSTRACT

Tests for comparing the strength of association between a variable X_1 and each of two potential predictor variables X_2 and X_3 are proposed and examined in a simulation study. The variances of X_2 and X_3 and the correlation between X_2 and X_3 are nuisance parameters. A simple modification of a test proposed by Williams (1959) is found to have good properties for a wide range of parameter values and both normal and nonnormal distributions.

INTRODUCTION

In a number of statistical settings, particularly in regression, it is desirable to know which of two random variables, say X_2 and X_3, is more strongly correlated with a dependent random variable X_1. Under the assumption that the observations are from a trivariate normal distribution, a number of tests for the hypothesis H_0: $\rho_{12} = \rho_{13}$ have been proposed. These have been analyzed and compared in some detail by Neill and Dunn (1975).

Further proposals have been made for a much more general setting where the underlying distribution cannot be regarded as normal and where the measure of strength of the relationship between the dependent and independent variables may be different than the Pearson product moment correlation coefficient. Hubert and Golledge (1981)

also discuss the situation where no specific population model is obvious.

Our intent is to examine a number of such suggestions, compare them with the procedures recommended in Neill and Dunn for the trivariate normal situation, observe their behavior under nonnormal distributions, and draw conclusions about their relative merits.

HISTORY

Let X_1, X_2, X_3 have a continuous trivariate distribution with covariance matrix Σ. Let $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ be the $(ij)^{th}$ element of Σ, with $\rho_{ii} = 1 (i = 1,2,3)$. For the trivariate normal, proposals for testing H_0: $\rho_{12} = \rho_{13}$ have been available since 1940, when Hotelling proposed as a test statistic the difference $r_{12} - r_{13}$ (where r_{ij} is the appropriate sample correlation coefficient) divided by an estimate of the asymptotic standard derivation of $r_{12} - r_{13}$. Neill and Dunn (1975) use both analytic methods and simulations in comparing eleven different test statistics including Hotelling's for this particular situation and recommend a statistic proposed by Williams (1959) as the best choice for small to moderate sample sizes.

Williams' test statistic, which also relies on a standardized version of $r_{12} - r_{13}$ and is only slightly different than Hotelling's proposal, is given by

$$T = (r_{12} - r_{13}) \sqrt{\frac{(n-1)(1 + r_{23})}{2(\frac{n-1}{n-3}) \cdot |R| + \bar{r}^2 (1-r_{23})}} ,$$

where

$$R = \begin{pmatrix} 1 & r_{12} & r_{13} \\ r_{12} & 1 & r_{23} \\ r_{13} & r_{23} & 1 \end{pmatrix}$$

and $\bar{r} = \dfrac{r_{12} + r_{13}}{2}$.

The one-tailed test compares T to the upper α^{th} percentile of the t-distribution on n-3 degrees of freedom.

As discussed in Boyer and Schucany (1978) a distribution free approach to this problem relies on observations by Wolfe (1976) that the correlation between X_1 and $Z = X_2 - X_3$ is given by

$$\rho_{1Z} = \frac{\rho_{12}\sigma_1\sigma_2 - \rho_{13}\sigma_1\sigma_3}{\sigma_1(\sigma_2^2 + \sigma_3^2 - 2\rho_{23}\sigma_2\sigma_3)^{1/2}}$$

and thus, if $\sigma_2 = \sigma_3$, then $\rho_{1Z} = 0$ if and only if $\rho_{12} = \rho_{13}$. The restriction that X_2 and X_3 have the same scale is also needed for Kendall's $T_{1Z} > 0$ to imply $\rho_{12} > \rho_{13}$.

A number of proposals for test statistics make use of this requirement by replacing sample values of X_2 and X_3 with a set of scores that will circumvent the scale problem. One possibility is to replace each of the elements of the $\underline{X}_2$ and $\underline{X}_3$ vectors by their integer ranks, $R(X_{2i})$ and $R(X_{3i})$ respectively. A problem that arises here is that $Z_i' = R(X_{2i}) - R(X_{3i})$ may well involve a substantial number of tied values. An additional possibility is replacing X_{2i} and X_{3i} by their expected normal scores, $N(X_{2i})$ and $N(X_{3i})$. This will reduce the magnitude of the tie problem but will still eliminate the scale difficulties. Then, any of the usual nonparametric measures of correlation could be used to

detect association between the X_{1i} and $Z_i^* = N(X_{2i}) - N(X_{3i})$. Note that there does not appear to be a familar population quantity corresponding to the relationship between X_1 and Z^*; nevertheless the technique does allow one to make general inferences about the relationships of X_1 to X_2 and X_3.

A second class of procedures that could be used to test the hypothesis of interest would involve transformation of the observations for each of X_1, X_2 and X_3 so that they are somewhat normal and then apply one of the normal theory methods (probably Williams' test) to the transformed data. It is felt that in nonnormal situations this procedure will yield a test statistic that is more stable than just using Williams' test on the raw data.

Several tests utilizing one of these methods or a simple extension of one of them provided a starting place for a substantial simulation study for comparison. A number of additional procedures, as described in Boyer and Schucany (1978) were also used in the early stages of the investigation, but proved to be inadequate even in the very simplest situation where the underlying distribution was trivariate normal and the null hypothesis H_0: $\rho_{12}=\rho_{13}$ was true. These methods were thus eliminated from subsequent parts of the study.

For instance, the procedure proposed by Davis and Quade (1968) uses Kendall's tau as the measure of correlation and a U-statistics approach to the hypothesis testing problem. However, the initial runs indicated that the empirical power was dominated by the Choi procedure. This combined with the additional fact that the U-statistic approach is more complicated computationally than the procedures using ranks, led to the procedure being dropped from the study.

THE SIMULATION STUDY

An extensive simulation study was run to compare the test statistics listed below. For each parameter configuration and distribution assumption 1000 samples of size 10 and 1000 samples of size 25 were generated and the appropriate one-tailed test performed at a nominal level of .05.

Using the IMSL subroutine GGNSM, the first samples were generated with X_1, X_2 and X_3 having the trivariate normal distribution with variance-covariance matrix

$$\Sigma = \begin{pmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{pmatrix}.$$

Note that the parameter ρ_{23} is a nuisance parameter which must be handled. Under H_0, the study used 53 parameter configurations which adequately cover all the possibilities for ρ_{23} and $\rho_{12} = \rho_{13}$ that give a positive definite covariance matrix. Under the alternative hypothesis, 52 different configurations, limited to the cases where both ρ_{12} and ρ_{13} are positive, were used. If the signs of ρ_{12} and ρ_{13} are known, as is often the case in practical situations, an appropriate change of sign on one of the variables can always be made so that ρ_{12} and ρ_{13} are positive. Some of the early runs included configurations where ρ_{13} or both parameters were negative, but all the results were strictly consistent with the case where both parameters are positive. So those situations were not used in subsequent runs.

Additional samples were generated under a trivariate lognormal distribution. Each observation was obtained by generating a trivariate normal observation (Z_1,Z_2,Z_3) and making the transformation $X_i = \exp(Z_i)$, $i = 1,2,3$. In order to obtain the covariance matrix Σ' for (X_1,X_2,X_3), the generating trivariate normal distribution has a covariance matrix with elements

$$\rho_{ij} = \log[\rho'_{ij}(e-1) + 1] ,$$

where the ρ'_{ij} are the desired elements of Σ'. This trivariate lognormal distribution not only has the advantage of being easy to generate, but it also has marginal distributions that are quite nonnormal.

There are fewer parameter configurations which give a positive definite covariance matrix for both this lognormal and the generating trivariate normal distribution, however. In the present study 30 such configurations which correspond to H_0 are reported, and 45 which fall in the region of the alternative was studied.

Five test statistics were evaluated in the full study (although, as mentioned previously, some early parts of the study included others). The five, with the abbreviations used in the tables of results are:

(W) Williams' test, as applied to the raw data. This is the benchmark, at least as far as the normal distribution is concerned, although its behavior under nonnormal circumstances had not been studied.

(C) The test proposed by Choi (1977). This requires replacing X_{2i}, X_{3i} by their respective ranks, $R(X_{2i})$

and $R(X_{3i})$, defining $Z_i' = R(X_{2i}) - R(X_{3i})$, computing the Spearman rank order correlation coefficient $r_s(X_1, Z')$ and comparing to the usual critical points for the Spearman coefficient.

(NS) The normal scores procedure proposed by Boyer and Schucany (1978). Exactly as (C) above except that the expected normal scores $N(X_{2i})$ and $N(X_{3i})$ are used in place of the ranks.

(WNS) Williams' test applied to the normal scores. That is, X_{1i}, X_{2i}, X_{3i} are replaced by $N(X_{1i})$, $N(X_{2i})$, $N(X_{3i})$ and then Williams' test is applied.

(WR) Williams' test applied to the ranks.

RESULTS AND CONCLUSIONS

Tables 1 and 2 present the results of the simulation study under the trivariate normality assumption and at parameter configurations consistent with H_0: $\rho_{12}=\rho_{13}$ for samples of size 10 and 25, respectively. The .05 level used here would imply that the particular test being considered ought to reject H_0 approximately 50 times, at any of these null parameter values.

The most readily apparent observations from the tables are that, as expected Williams' test very consistently rejects H_0 about 5% of the time, and that both C and NS tend to be extremely conservative when the magnitude of ρ_{12} and ρ_{13} is large (in fact when $\rho_{12}=\rho_{13}$ = .9, neither test rejected any of 1000 samples of size 25 and together they rejected only 3 times for samples of size 10). It is clear, in fact, that these two tests, which are based on rank correlation and thus might be expected to be distribution

free, are not even parameter free. Note also that the two procedures which replaced the data by scores (either ranks or normal scores) and then used Williams' procedure behaved well. WR rejected 29 and 91 times in the two most extreme cases while WNS rejected 32 and 88 times in its most extreme cases. Although neither is as stable as Williams' test (as expected), likewise neither suffered any serious difficulties in maintaining something close to the nominal level for the test.

The power study at the normal distribution tends to confirm the suppositions in the preceding paragraphs. Since $\rho_{12} \neq \rho_{13}$ causes the parameter space to be three-dimensional, the entire study does not lend itself readily to tables. However, we illustrate the points with a few sequences of parameters chosen from the study with ρ_{13} and ρ_{23} fixed and ρ_{12} moving away from ρ_{13}, a case in which we expect to see increasing power. Three such examples appear as Table 3. In each case we see that C and NS have considerably less power than the competing procedures. (In at least one case that observation must be tempered by noting that C and NS fell significantly below the nominal level at the null hypothesis, and thus might be expected to fall short in the power comparisons at nearby parameter configurations as well.) We also note again that while WNS and WR do not achieve the same power as Williams' test, they do not fall disastrously short of the desired performance.

Tables 4 and 5 illustrate the performance of the same test statistics at the null hypothesis when the trivariate distribution

has lognormal marginals. Several important observations need to be made here. First, as before C and NS do not maintain the desired .05 level. Again the parameter configuration where they had the most difficulty were those which had the greatest magnitude for ρ_{12} and ρ_{13}, and, as before, they tend to be extremely conservative at those values.

Second, as might have been suspected, the behavior of Williams' test breaks down for this highly skewed distribution. For sample size 10, the observed significance level varies from .007 to .149 with 20 of the 30 parameter configurations giving values outside the interval .037 to .063 (which is .050 $\pm$ 2 standard errors) and for samples of size 25, the observed significance level varies from .002 to .214 with 24 of the 30 parameter configurations giving values outside the .037, .063 interval. On the other hand, the tests that replace the data by scores and then apply Williams' procedure fared much better. WNS was outside the interval .037 to .063 only 3 of 30 times for samples of size 10 and 2 of 30 times for samples of size 25, while the figures are 6 of 30 at $n = 10$ and 5 of 30 at $n = 25$ for the WR procedure.

In Table 6, sample sequences of parameter configurations which move away from H_0 are again considered. It should be noted here that the C and NS procedures have lower power than the other procedures in general. It should also be noted that in the last example the procedures using the score functions surpass Williams' test in terms of power as the ρ_{12} and ρ_{13} become more separated, even though Williams' procedure had a large observed significance level at H_0.

The results here are typical of those of the whole study in that, in most cases, the WR and WNS procedures were competitive with W, never having an inordinately smaller power. In fact, in some cases where W has more power, it appears attributable to the fact that the true level of W is not very stable for this distribution.

RECOMMENDATIONS

In situations where a practitioner is comfortable with the assumption of trivariate normality, it is recommended that Williams' test be used. This is consistent with Neill and Dunn (1975). On the other hand, when normality is not a good assumption it is recommended that WR or WNS be used, as their behavior is much more stable than Williams' test, and competitive in terms of power. Between the two tests, the choice might be difficult. Using the power study, WNS appears to be slightly better. On the other hand, use of the ranks does not require special tables and it appears that the computation, particularly if it is to be done by hand, might be sufficient to recommend the WR procedure.

In retrospect, one notices that replacing the data by ranks and then applying the usual normal theory techniques to make the appropriate inference is an idea that Iman and Conover (see Iman (1974) or Iman and Conover (1979)) have espoused in a number of other statistical settings.

REFERENCES

Boyer, J.E. and Schucany, W.R. On Wolfe's test for related correlation coefficients. SMU Technical Report No. 127, 1978.

Choi, S.C. Tests of equality of dependent correlation coefficients. Biometrika, 1977, 64, 645-647.

Davis, C. E. and Quade, D. On comparing the correlation within two pairs of variables. Biometrics, 1968, 24, 987-995.

Hotelling, H. The selection of variates for use in prediction with some comments on the problem of nuisance parameters. Annals of Mathematical Statistics, 1940, 11, 271-83.

Hubert, L. J. and Golledge, R. G. A heuristic method for the comparison of related structures. Journal of Mathematical Psychology, 1981, 23, 214-26.

Iman, R. L. A power study of a rank transform for the two way classification model when interaction may be present. Canadian Journal of Statistics Applications, 1974, 2, 227-39.

Iman, R. L. and Conover, W. J. The use of the rank transform in regression. Technometrics, 1979, 21, 499-510.

Neill, J. J. and Dunn, O. J. Equality of dependent correlation coefficients. Biometrics, 1975, 31, 531-543.

Williams, E. J. The comparison of regression variables. Journal of the Royal Statisticial Society, Series B, 1959, 21, 396-399.

Wolfe, D. A. On testing equality of related correlation coefficients. Biometrika, 1976, 63, 214-215.

TABLE 1

Monte Carlo estimates of True Significance Levels

(Number of times H_0 rejected in 1000 trials)

Normal Distribution, n=10, Nominal α=.05

ρ_{23}		$\rho_{12}=\rho_{13}$: -.9	-.7	-.5	-.3	-.1	0	.1	.3	.5	.7	.9
	W	59	58	66	45	55	45	45	53	53	37	49
	C	3	21	36	42	39	47	45	39	29	15	1
.9	NS	3	23	41	45	46	62	57	51	32	14	2
	WNS	88	61	58	50	53	55	48	53	46	40	75
	WR	85	63	45	55	53	47	59	47	59	59	91
	W		42	34	54	38	43	44	36	59	36	
	C		4	29	39	39	44	40	38	33	10	
.6	NS		7	25	47	48	49	47	44	38	18	
	WNS		55	43	48	47	48	44	49	49	55	
	WR		50	55	46	51	53	50	45	60	59	
	W		45	45	42	45	43	42	43	47	48	
	C		7	27	46	48	40	41	35	27	7	
.3	NS		8	25	46	53	45	44	37	26	9	
	WNS		55	50	42	52	39	41	41	46	45	
	WR		72	68	42	50	40	57	52	51	70	
	W		45	45	38	39	42	51	42	38	41	
	C		3	17	31	37	45	45	31	24	7	
0	NS		5	22	33	41	46	46	42	28	8	
	WNS		41	51	42	33	42	47	43	50	57	
	WR		54	52	50	42	52	57	59	45	37	
	W			39	40	49	40	48	46	37		
	C			15	31	55	24	40	36	20		
-.3	NS			17	36	59	31	47	45	17		
	WNS			35	42	44	30	41	43	37		
	WR			59	45	53	51	54	42	43		
	W				33	38	33	33	45			
	C				27	33	48	35	35			
-.6	NS				31	38	47	39	33			
	WNS				32	36	38	37	37			
	WR				40	49	34	55	44			
	W					33	40	36				
	C					44	54	42				
-.9	NS					48	59	46				
	WNS					38	41	34				
	WR					36	46	34				

TABLE 2

Monte Carlo estimates of True Significance Levels

(Number of times H_0 rejected in 1000 trials)

Normal Distribution, n=25, Nominal α = .05

$\rho_{12} = \rho_{13}$

ρ_{23}		-.9	-.7	-.5	-.3	-.1	0	.1	.3	.5	.7	.9
	W	53	54	58	60	50	48	51	48	45	36	29
	C	3	21	37	43	55	41	41	35	24	8	0
.9	NS	4	23	37	51	58	44	51	36	27	8	0
	WNS	88	57	43	39	53	42	57	47	45	47	53
	WR	66	57	47	53	42	49	49	49	48	64	81
	W		48	50	56	50	48	49	46	47	37	
	C		13	26	43	49	43	40	42	22	8	
.6	NS		13	32	47	49	52	43	42	20	4	
	WNS		54	51	46	58	60	46	43	43	45	
	WR		62	67	46	47	50	67	52	50	43	
	W		51	40	44	45	45	57	55	57	48	
	C		4	23	38	43	42	46	39	28	9	
.3	NS		4	24	40	53	48	54	48	26	10	
	WNS		55	39	44	49	48	55	55	48	49	
	WR		50	64	45	49	55	38	47	38	66	
	W		48	52	54	67	60	55	38	44	54	
	C		1	26	45	49	51	46	38	22	4	
0	NS		1	26	46	52	57	54	40	28	6	
	WNS		44	53	56	63	59	46	41	54	59	
	WR		58	45	55	43	45	41	58	51	62	
	W			39	43	43	44	37	41	47		
	C			10	29	50	44	33	25	20		
-.3	NS			12	27	53	54	45	32	22		
	WNS			37	45	44	44	47	42	46		
	WR			60	51	42	43	40	46	46		
	W				49	42	38	35	44			
	C				40	32	42	45	33			
-.6	NS				43	40	46	47	42			
	WNS				48	40	41	45	46			
	WR				58	33	44	47	47			
	W					40	39	46				
	C					39	42	43				
-.9	NS					45	44	44				
	WNS					42	47	50				
	WR					43	54	29				

TABLE 3

Monte Carlo estimates of Power of the Tests
(Number of times H_0 rejected in 1000 trials)

Normal distribution, Nominal $\alpha = .05$

$\rho_{23} = .9$, $\rho_{13} = .5$, $n = 10$

	ρ_{12}		
	.5	.6	.8
W	53	182	982
C	29	77	410
NS	32	89	425
WNS	46	131	727
WR	59	127	758

$\rho_{23}=0$, $\rho_{13}=.1$, $n = 25$

	ρ_{12}				
	.1	.2	.4	.6	.8
W	55	78	255	666	953
C	46	81	196	499	817
NS	54	84	221	524	822
WNS	46	80	239	623	930
WR	41	90	280	580	924

$\rho_{23}=-.3$, $\rho_{13} = .3$, $n = 25$

	ρ_{12}			
	.3	.4	.6	.8
W	41	97	299	797
C	25	58	165	355
NS	32	70	169	376
WNS	42	90	282	717
WR	46	91	281	689

TABLE 4

Monte Carlo estimates of True Significance Levels
(Number of times H_0 rejected in 1000 trials)
Lognormal Distribution, n=10, Nominal α = .05

$\rho_{12} = \rho_{13}$

ρ_{23}		-.3	-.1	0	.1	.3	.5	.7	.9
	W	20	42	48	63	95	109	131	149
	C	6	45	44	33	30	20	7	1
.9	NS	11	49	54	48	33	22	7	3
	WNS	56	45	47	44	47	45	53	98
	WR	58	47	50	44	64	44	61	112
	W	14	41	54	55	101	128	129	
	C	12	35	40	37	33	13	4	
.6	NS	16	47	47	49	37	17	6	
	WNS	70	47	45	47	41	41	60	
	WR	68	43	47	45	44	66	64	
	W	7	32	53	70	96	116	141	
	C	4	41	42	47	26	15	3	
.3	NS	8	55	45	49	21	19	5	
	WNS	42	61	46	44	39	53	59	
	WR	42	60	49	52	51	52	67	
	W		40	47	77	92	143		
	C		37	38	45	32	17		
0	NS		41	40	48	29	20		
	WNS		37	37	42	49	48		
	WR		40	48	41	47	58		
	W		32	56	70				
	C		41	47	32				
-.3	NS		42	51	38				
	WNS		44	42	35				
	WR		44	42	47				

TABLE 5

Monte Carlo estimates of True Significance Levels
(Number of times H_0 rejected in 1000 trials)
Lognormal Distribution, n = 25, Nominal α=.05

ρ_{23}		$\rho_{12}=\rho_{13}$: -.3	-.1	0	.1	.3	.5	.7	.9
	W	30	29	54	75	111	133	147	195
	C	11	43	41	38	26	20	10	0
.9	NS	18	45	45	52	28	23	11	0
	WNS	55	49	55	51	45	46	51	72
	WR	56	63	39	56	50	50	61	75
	W	15	35	49	69	113	133	199	
	C	10	45	41	46	34	24	5	
.6	NS	10	47	47	50	29	27	5	
	WNS	47	51	53	57	57	57	56	
	WR	57	40	56	64	56	61	64	
	W	2	40	51	72	121	160	214	
	C	6	54	45	42	31	18	4	
.3	NS	8	60	49	50	37	19	5	
	WNS	38	65	50	50	47	51	49	
	WR	64	55	43	58	38	56	78	
	W		35	54	81	134	173		
	C		51	38	38	33	8		
0	NS		56	41	48	32	7		
	WNS		62	54	56	53	47		
	WR		46	47	42	47	54		
	W		26	48	69				
	C		48	41	40				
-.3	NS		53	43	41				
	WNS		59	45	48				
	WR		43	53	37				

TABLE 6

Monte Carlo estimates of Power of the Tests
(Number of times H_0 rejected in 1000 trials)
Lognormal distribution, Nominal $\alpha = .05$

$\rho_{23}=.6$, $\rho_{13}=.1$, $n = 10$

	ρ_{12} .1	.2	.4	.6
W	55	132	348	778
C	37	88	249	607
NS	49	101	266	605
WNS	47	104	314	757
WR	45	109	363	725

$\rho_{23}=.3$, $\rho_{13}=.3$, $n = 10$

	ρ_{12} .3	.4	.6	.8
W	96	153	319	578
C	26	50	113	191
NS	21	63	121	207
WNS	39	80	204	497
WR	51	88	222	485

$\rho_{23}=0$, $\rho_{13}=.1$, $n = 25$

	ρ_{12} .1	.2	.4	.6	.8
W	81	142	355	642	918
C	38	91	274	568	843
NS	48	104	294	592	858
WNS	56	105	366	743	979
WR	42	112	380	716	971

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: 158
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:
4. TITLE (and Subtitle): Related Correlation Coefficients
5. TYPE OF REPORT & PERIOD COVERED: TECHNICAL REPORT
6. PERFORMING ORG. REPORT NUMBER: 158
7. AUTHOR(s): John E. Boyer, Albert D. Palachek and William R. Schucany
8. CONTRACT OR GRANT NUMBER(s): N00014-82-K-0207
9. PERFORMING ORGANIZATION NAME AND ADDRESS: Southern Methodist University, Dallas, Texas 75275
10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS: NR 042 479
11. CONTROLLING OFFICE NAME AND ADDRESS: Office of Naval Research, Arlington, Va. 22217
12. REPORT DATE: March 1982
13. NUMBER OF PAGES: 17
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office):
15. SECURITY CLASS. (of this report):

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report): This document has been approved for public release and sale; its distribution is unlimited. Reproduction in whole or in part is permitted for any purposes of the United States Government
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report):
18. SUPPLEMENTARY NOTES:
19. KEY WORDS (Continue on reverse side if necessary and identify by block number): Correlation coefficient, normal scores hypothesis testing, approximately distribution-free

20. ABSTRACT (Continue on reverse side if necessary and identify by block number) Tests for comparing the strength of association between a variable X_1 and each of two potential predictor variables X_2 and X_3 are proposed and examined in a simulation study. The variances of X_2 and X_3 and the correlation between X_2 and X_3 are nuisance parameters. A simple modification of a test proposed by Williams (1959) is found to have good properties for a wide range of parameter values and both normal and nonnormal distributions.